이진집합기반에서 칼라와 형태정보를 이용한 영상 검색시스템 설계

김성동[†] · 최기호[††]

요 약

본 논문에서는 이진 집합하 에서 칼라 공간과 형태 정보를 가지고 새로운 영상검색 방법을 제시한다. 각 영상에 대한 칼라 공간 정보는 칼라 이진세트에 의해 구해지고 형태정보는 영역 세그멘테이션에 의해서 구해진다. 영상 검색 과정에서, 질의 영상과 데이터베이스 영상들의 칼라 및 영상 이진세트들을 비교하여 검색될 후보 영상의 집합이 결정된다. 특히, 유사도 측정은 시각적으로 유사한 칼라들의 분포와 객체의 칼라공간 및 형태 특징에 가중치를 고려한 검색이 가능하도록 하였다. 또한 검색 과정의 속도를 향상시키기 위해 후보영상에 대한 복잡한 유사도 측정을 단순히 수행할 수 있도록 새로운 검색 기법을 제안하였다. 위에서 제안한 방법과 3,000개의 화상들로 이루어진 영상 데이터베이스에 대한 구현 실험을 수행하여 제안된 칼라 공간 및 형태특징을 기반으로 한 영상 검색이 비교적 효과적인 결과를 보였다.

Design of Image Retrieval System using Color and Morphological Informations based on Binary Sets

Seong Dong, Kim[†] and Ki Ho, Choi[††]

ABSTRACT

This paper presents a new image retrieval system with color and morphological informations based on binary sets. Each of them can be obtained from color binary sets and regional segmentation separately. For retrieval processes, the candidate images are decided by comparing color and their image binary sets of the database with query images. Particularly, it is possible that the retrieval of similar-measurements has a weight of color spatial distribution and its objective morphological features. We proposed a new idea for performing simply the complicated similar-measurement of candidated images to improve queried processes. The retrieval method using spatial and morphological features is shown with the effectiveness on the result of implementation on database with 3,000 images.

1. 서 론

몇 년동안 영상검색에 대한 다양한 연구방법이 제시되어 왔다. 이러한 시스템에서의 검색 과정은 객체의 칼라, 구조, 형태 등 화상의 가시적인 속성이나 그림 요소를 근거로 하고 있다. 초기에 데이터베이스 화상의 내용(contents)은 (반)자동적으로 추출되고 저장되어 진다. 질의과정 동안 질의 영상의 내용이 자동추출되고 이 추출된 내용은 데이터베이스에 있는 여러 영상들과 매칭하고, 검색 결과 질의영상과의 정확한 매칭이 아닌 유사도 매칭한 매칭방법에 의해 영상들이 출력된다.

또한 칼라는 영상을 표현하는데 광범위하게 사용되어 왔다[5,7,8,11,14,18,19]. 일반적으로 밝기 값의 히스토그램으로 표현되는 영상의 칼라 구성은 영상을 이루고 있는 객체들에 대한 정보를 포함하지 않는 전역적인 속성이다. 더군다나, 칼라 분포는 시각과 해상도에 독립되고, 칼라비교는 인간의 간섭없이 자

[†] 정회원, 계원조형예술대학 정보통신과 교수
[††] 종신회원, 광운대학교 컴퓨터공학과 교수

동적으로 실행될 수 있다. 그러나, 칼라만으로 영상을 특징화하기에는 부족한데, 그 이유는 칼라들이 군집되어 있는 방법과 위치가 서로 다르기 때문이다.

그 동안 보다 정확한 영상검색을 위해 칼라 공간 분포와 칼라 정보를 통합한 칼라-공간 검색 기술이 여러 논문에서 연구되어 왔다[7,8,11]. 그러나 이것들은 계산시간을 소모하고 저장시 비능률적이었다. 또한, 상표나 마크같은 단순한 객체를 가지는 영상에 대한 모양을 이용한 검색연구가 있었으며, 이는 복잡한 영상은 적용할 수 없다는 제약이 있다. 또한 기존 연구들에서는 칼라-공간과 형태 검색기술을 통합한 시스템이 없었다.

2. 관련 연구

그동안 텍스트 기반, 내용 기반[1,3,11,12,14,10,16, 19] 그리고 의미 기반[4]과 같은 영상 검색 기술에 대한 많은 연구가 있었다. [11]에서는 쿼드 트리구조에 기반한 다중레벨 칼라 히스토그램 접근 방법이 칼라의 위치 정보를 다루는데 사용되어졌다. Top-level 칼라 히스토그램은 전체 영상의 칼라 구성을 표현하고, 다음 단계는 4등분한 영상의 칼라 분포에 해당하는 네 개의 히스토그램으로 이루어져 있다. 일반적으로 i^{th}번째 레벨에서는 영상은 4^{i-1}의 영역으로 나누어지며 각 영역은 칼라 구성에 대한 히스토그램을 가지고 있다. 질의를 하는 동안 원 영상과 목적 영상들은 히스토그램의 칼라 구성을 기반으로 하여 비교되어지는데 맨 위 레벨의 히스토그램이 처음 비교되어진다. 만약 맨 위 레벨에서 서로 매치 된다면 그 다음 레벨이 탐색될 것이다. 단지 그 레벨에서의 임계치와 일치할 때 만 영상이 링크되어지고 검색되어 진다. 그러나 이러한 방법은 컴퓨터의 저장 측면에서 효율적이지 못하다.

[7]에서는 위에서 제시한 방법을 응용하였다. 전체 영상에 대한 대한 전역 히스토그램과 더불어 영상을 9개의 부분으로 나누어 각 영역에 대한 지역 히스토그램을 가지고 있다. 검색 과정 동안 사용자는 단지 전역 히스토그램이나 9개의 지역 히스토그램 전부 혹은 이들들을 조합하여 매칭할 수 있도록 선택할 수 있다. 유사도 계산을 위해 히스토그램 인터섹션(intersection)을 사용하는 대신에 각 binary set에 유일한 숫자 키를 부여함으로서 복잡한 히스토그램 매칭 과정을 간단한 숫자 키로서의 비교로 바꾸어 준다.

위의 2가지 방법에서는 칼라들의 공간 분포가 다양한 히스토그램에 의해 함축적으로 추출된다. [8]에 의해 제시된 방법은 보다 나은 기법으로 단일-칼라 영역과 공간 지역성을 확인한다. 알고리즘은 두 가지 단계로 이루어져 있는데 첫 번째 단계에서는 영상의 대표칼라 집합이 선택되고 두 번째 단계에서는 선택된 칼라들 각각에 대한 클러스터 집합들이 최대 엔트로피 분포 방법[2]에 의해 결정된다. 이러한 단일 칼라 클러스터들은 사각형으로 표현되어진다. 영상간의 유사도 측정은 원영상과 목적 영상에서의 동일 칼라의 직사각형들 사이의 중첩되는 양에 의해 결정되어진다. 만약 두 영상이 같은 칼라를 가져 2개의 직사각형이 중첩이 되면 영상의 공간에서 교차된다. 그리고 중첩의 정도는 교차되는 픽셀의 수로 주어진다. 비록 이 방법이 히스토그램을 기반으로 한 방법보다 우수한데 반해 컴퓨터로 유사도를 측정하는 기능과 하나의 색으로 구성된 지역을 추출하는데 있어 계산시간이 많이 걸린다. 더군다나 하나의 영상에서 다른 영상으로 변할 때 만들어지는 칼라 그룹의 수는 변하게 되고 매우 커질 수도 있어서 이는 검색 속도를 매우 느리게 만들 수도 있다.

3. 칼라-공간 및 형태 검색

본 절에서는 본 논문에서 제안한 이진세트 기반에서 칼라-공간 및 형태검색 기술로 부터 칼라-공간 및 형태 정보가 추출되고 표현되는 과정을 기술한다. 특히, 검색과정은 두 영상 사이의 칼라-공간 유사도를 계산하기 위한 기준이 필요하다. 그리고 시각적으로 유사한 칼라 개념과 칼라들의 가중치를 통합한 접근방법을 제안한다.

3.1 칼라-공간 정보의 표현

이러한 방법이 실용적이고 유용하게 하기 위해서 몇 가지 고려해야할 점은 첫째로 칼라의 수가 매우 많아질수록 결과적으로 칼라-공간 정보의 집합이 커진다. 이를 해결하기 위해 각각의 영상에 대해 칼라의 수를 영상의 대표 칼라 C의 집합으로 제한하면 된다. C는 대부분의 영상들이 몇 가지 칼라들에 의해 두드러지듯 적을 것으로 예상된다. 대표 칼라 C를 선택하기 위해 경험적인 방법을 채택하였다. H_i 칼

라 히스토그램은 영상의 칼라구성을 표현하고 있다. 여기서 H_i에서 가장 큰 픽셀 수를 가진 $C_i(C_i<C)$ 칼라를 선택하였다.

두 번째 문제는 칼라-공간 정보를 표현하는 것과 관계가 있다. 본 논문에서 제안한 방법은 주어진 칼라에 의해 셀이 표현되어질 수 있고 표현되지 않을 수도 있는 훌륭한 특성을 가지고 있다. 각 셀은 그 자체로서 비트에 의해 표현되어질 수 있는데 만약 셀이 임계치를 만족하면 비트는 세트(1)될 것이고 그렇지 않으면 클리어(0)될 것이다. 그러므로 각 칼라에 대해 비트스트림(bitstream)을 가지며 이를 칼라 이진세트(binary sets)라 한다.

예를 들어, 그림 1에서 칼라가 셀 0,1,2,5,6,7,8,9, 10,16,17,18,24,25,26,27,44,45,46,47,52,53,54,55,60,61, 62, 63에 의해 표현되어 진다고 가정하면 그것에 대응하는 64비트 칼라 이진세트는 1110011111100000 1110....1111 0000111100001110과 같은 비트 스트림이 될 것이다. 주어진 영상의 대표 칼라 C에 대해 칼라 이진세트를 가지게 된다. 그리고, 칼라 공간정보는 객체의 상하관계, 좌우관계를 나타나게 하였다.

다음 그림 2에 칼라공간 특징추출 과정을 간략히 나타내었다. 입력된 RGB 원영상은 HSV 칼러공간 변환을 걸쳐 32칼라 양자화와 필터링 과정을 거쳐서 8×8 영역으로 분할한 다음, 전역 칼라 히스토그램을 추출한다.

3.2 형태 정보의 표현

일반적인 세그멘테이션 과정은 그레이레벨 차이

			0	0			
			0	0	0	0	0
			0	0	0	0	0
				0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0				
0	0	0	0				
0	0	0	0				0
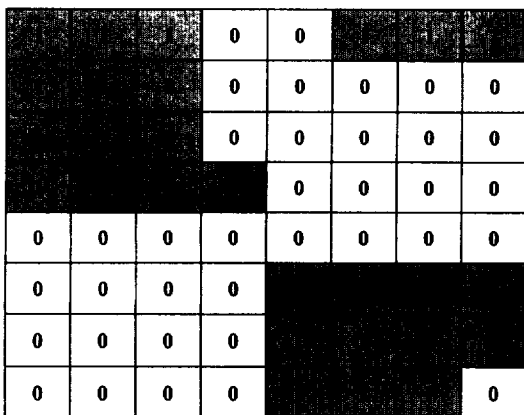
그림 1. 8×8로 분할된 원영상 예
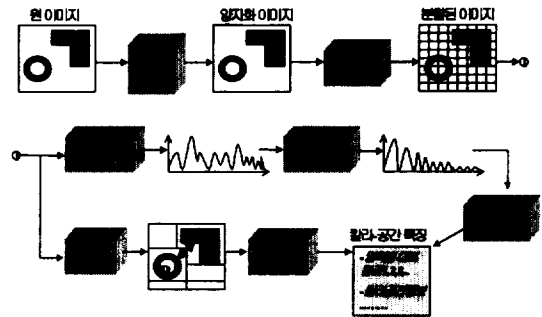
그림 2. 칼라공간특징 추출을 위한 세그멘테이션

를 기초로 이루어 지며, 현재 그레이레벨로 region split & merge하는 방법들도 여러 연구에서 많이 사용하고 있으며, 칼라 이미지 프로세싱이 쉬워지면서 칼라 차이에 의한 세그멘테이션 방법도 증가하고 있다. 현재 과학적인 이미지 프로세싱에서 사용되는 접근방법은 인간의 세그멘테이션 능력을 모방하려는 소위 "지식기반(knowledge-based) 세그멘테이션" 방법이다.

본 논문에서는 256그레이 레벨을 64 그레이레벨로 양자화 한 후 그레이 레벨에 따라 라벨이미지를 얻고 연결성분을 분석하여 마크이미지를 생성해 내어 모양특징을 추출하였으며 칼라와 텍스쳐 정보를 고려하지 않았기 때문에 세그멘테이션 결과는 칼라와 텍스처를 고려한 영역 세그멘테이션보다는 성능면에서 개선되었다고는 주장할 수 없다.

그림 3과 그림 4는 영역기반 세그멘테이션 과정과 형태추출과정을 각각 보이고 있다. 먼저, 원영상의 그레이 히스토그램을 생성하고 히스토그램 분석(histogram analysis)으로 시작한다. 본 예제 그림에서는 원영상이 배경과 객체를 표현하는 two 그레이레벨로 구성되어 있다고 가정한다. 히스토그램은 원
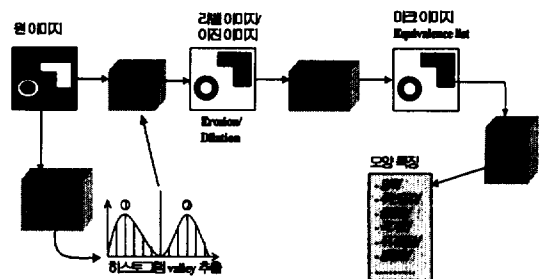
그림 3. 영역기반 세그멘테이션
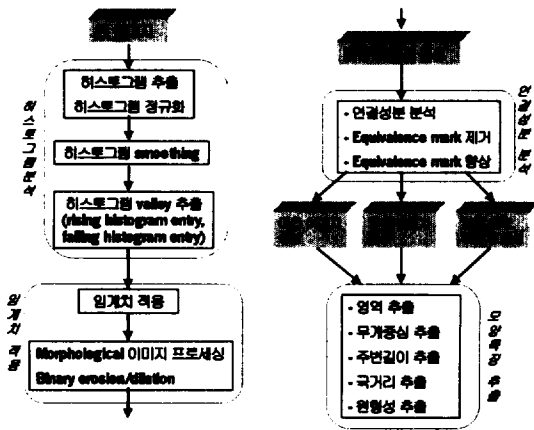
그림 4. 형태추출 과정

영상에 나타나는 그레이레벨의 빈도수를 나타내므로 이 경우에 히스토그램은 2개의 분리된 그레이레벨 영역으로 나누어지는 peak를 구성한다. 임계치(threshold) 적용에서는 두 히스토그램 peak간의 valley를 측정하여 valley 사이에 임계치를 할당하고 나서 이 임계치와 peak를 비교하여 임계치보다 작으면 라벨(label) '0'으로 설정하고 높으면 라벨 '1'로 설정하여 라벨 이미지(label image) 혹은 이진 이미지(binary image)를 얻는다.

원영상으로부터 분리된 영역들을 분석하기 위해서는 이미지의 모든 픽셀에 대한 정보를 추출해야 한다. 현재 영역기반 세그멘테이션에 의한 영역분석은 이들 영역에 대한 몇가지 전형적인 특징들에 기초하고 있으며, 본 논문에서는 다음과 같은 특징들을 추출하였다.

- 영역(area) : 영역의 크기, 면적
- 주변길이(perimeter) : 영역의 둘레, 길이
- 원형성(circularity, compactness) : $perimeter^2/(4\pi \times area)$로서 영역의 둥근정도
- 극거리(Polar distance) : distance-versus-angle signature로서 영역의 무게중심과 경계들간 거리
- 무게중심(center of gravitiy) : 영역의 위치

원의 경우에는 원형성은 1이다. 이것은 영역의 주변길이가 영역크기보다 길어지면 증가한다. 극거리는 영역의 무게중심과 영역의 경계간의 거리를 나타낸다. 원의 경우 극거리는 경계상의 어떤 점에 대해서도 동일하며 다른 모든 모양들은 경계상의 각 점마다 다양한 거리를 갖는다. 이러한 변화 형태가 영역모양을 특징화할 수 있다. 대부분의 특징들은 영역의 위치, 회전, 확대축소에 종속되므로 상황에 따라 유용할 수도 불필요할 수도 있다. 예를 들어, 무게중심은 영역의 위치에 종속되므로 영역의 위치를 결정할 때는 매우 유용하게 사용될 수 있다. 원형성은 비율측정이므로 위치, 회전, 확대축소에 독립적이며, 단순한 모양특징으로 유용하게 사용할 수 있다.

3.3 검색 유사도 측정

인간의 시각적인 관점에서는 만약 두 영상의 칼라 구성과 칼라 분포가 유사하면 두 영상은 똑같이 인식되어진다. 본 논문에서 제안한 칼라 이진세트는 영상을 대표 칼라의 집합으로 표현되므로, 두 영상을 표현하는 이진세트는 비록 몇몇 비트들만 다를지라도 비슷하다. 따라서, 부분적인 칼라에 대해 두 영상의 공통 부분을 계산하기 위해 간단한 논리적 AND 연산을 사용한다. 다음 그림 5는 입력영상에 대하여 칼라와 모양, 공간정보를 매칭하는 과정을 보이고 있다. 유사한 칼라와 모양을 갖는 영역 매칭후 최종적으로 공간관계를 갖는 영상 매칭을 수행하게 된다.
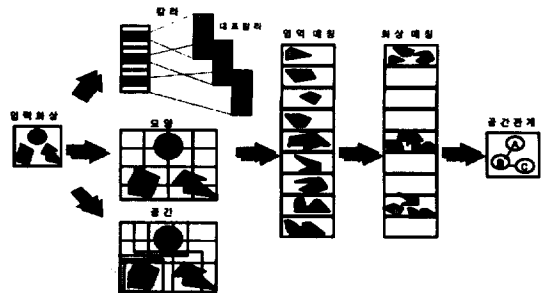
그림 5. 칼라, 형태, 공간 특징을 고려한 검색기법

3.3.1 칼라 유사도 계산 함수

본 논문의 칼라 검색 접근방법을 위해 영상에서 특정셀을 나타내는 이진세트에서의 각 비트를 고려할 수 있다. 질의 영상 Q와 데이터베이스 영상 D의 각각의 칼라 i의 이진세트를 C_{Qi}와 C_{Di}으로 나타내면 두 이진세트에서 일치하는 비트가 세트될 경우 두 영상은 똑같은 영역 즉, 셀에서 칼라 i를 가지게 되며 만약 그렇지 않을 때는 그 영역에서는 두 영상이 유사하지 않다는 것을 뜻한다. 그리고 Q와 D의 칼라 영

역 유사도를 측정하는데 있어 칼라 i가 C_Q에 포함될 때 Q와 D사이의 S_b의 값은 다음식과 같이 결정된다.

$$S_b(Q,D,i)=\begin{cases}\dfrac{BitSet(C_{Qi}\wedge C_{Di})}{BitSet(C_{Qi})} & \text{if } i\in C_{Di}\\ 0 & otherwise\end{cases} \quad (1)$$

여기서 BitSet(BS)은 비트 스트림 BS에서의 세트되는 비트들의 수를 나타내며 '∧'은 논리 AND 연산을 나타낸다. 만약 Q의 넓은 부분의 셀들이 D와 똑같은 칼라를 가진다면 계산되어지는 유사도값은 1이 될 것이다. 두 영상 Q 와 D사이의 칼라 유사도 CS_b의 측정은 다음식과 같게 된다.

$$CS_b(Q,D)=\sum_{\forall i\in C_Q} S_b(Q,D,i) \quad (2)$$

그리고 $CS_b(Q, D)$는 0과 1사이의 값으로 미리 정해진 임계치와 비교된다. 임계치는 두 칼라간의 유사도가 고려되어 질 때의 오차 허용도를 나타낸다. 만약 $CS_b(Q, D) > 0$ 이면 칼라 Q는 칼라 D와 시각적으로 유사한 정도를 나타낸다. $CS_b(Q, D)$값이 더 커질수록 두 칼라는 더 유사하고, 만약 $CS_b(Q, D)=0$이면 두 칼라간에는 유사도가 없다는 것을 뜻한다. 모든 칼라 이진세트 쌍들에 대해 계산되어진 유사값들은 similarity matrix(SM)이라 불리는 M×N행렬에 저장된다. 여기서, 대표칼라수 M = 8이고 N = 1을 나타낸다.

3.3.2 공간 유사도 계산 함수

본 논문의 공간 검색 접근방법을 위해 질의 영상 Q와 데이터베이스 영상 D의 각각의 영역 i의 공간위치 S_{Qi}와 S_{Di}으로 나타내면 두 영역위치간의 공간 유사도를 측정하는데 있어 영역 i를 S_Q에 포함할 때 Q와 D사이의 SS_b의 값은 다음과 같이 결정된다.

그림 6에서는 질의영상과 참조영상간의 임의 한
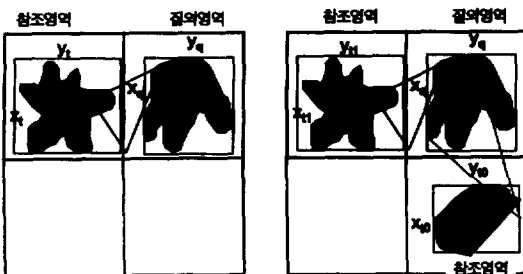
그림 6. 공간 특징의 유사도 측정방법

영역의 공간위치를 비교하는 것으로 두 영역이 겹쳐지는 정도에 따라 영역이 유사한지 아닌지를 결정하며 질의영상 Q의 영역위치를 (x_q, y_q)라 하고 데이터베이스영상 D의 영역위치를 (x_t, y_t)로 나타낸다. 따라서, Q와 D의 공간 유사도를 측정하는데 있어 공간 영역 i를 S_Q에 포함할 때 Q와 D사이의 SS_b의 값은 다음식과 같이 결정된다.

$$SS_b(Q,D)=\sum_{\forall i\in S_Q}\sqrt{(x_{qi}-x_{ti})^2+(y_{qi}-y_{ti})^2} \quad (3)$$

$SS_b(Q, D)$값이 작을 수록 두 영역은 더 유사하고, 만약 $SS_b(Q, D)$이 크면 두 영역간에는 유사도가 없다는 것을 뜻한다. 모든 영역의 쌍들에 대해 계산되어진 유사값들은 similarity matrix(SM)이라 불리는 M×N 행렬에 저장된다.

3.3.3 형태 유사도 계산 함수

본 논문의 형태 검색 접근방법은 Euclidean 거리로 두 형태 정보들간의 유사도를 측정할 수 있다. 여기서, 이미지의 형태정보들은 각각 영역 Ma, 주변길이 Mp, 원형성 Mc, 극거리 Md, 무게중심 Mg들이 되고 이들 형태정보 i와 j사이의 Euclidean 거리를 다음식과 같이 계산할 수 있다.

영역 유사도 : $M_{aD}(i,j)=\sqrt{(M_{ai}-M_{aj})^2}$ (4a)

주변길이 유사도 : $M_{pD}(i,j)=\sqrt{(M_{pi}-M_{pj})^2}$ (4b)

원형성 유사도 : $M_{cD}(i,j)=\sqrt{(M_{ci}-M_{cj})^2}$ (4c)

극거리 유사도 : $M_{dD}(i,j)=\sqrt{(M_{di}-M_{dj})^2}$ (4d)

무게중심 유사도 : $M_{gD}(i,j)=\sqrt{(M_{gi}-M_{gj})^2}$ (4e)

위 식은 형태정보에 대한 유사도를 표현하는 것이며 형태 정보들의 각각 유사도의 거리 M_{aD}, M_{pD}, M_{cD}, M_{dD}, M_{gD}를 나나태며, 이를 질의영상 Q와 데이터베이스영상 D의 형태정보 세트 유사도 M_{Qi}와 M_{Di}으로 나타낸다. 따라서, Q와 D의 형태 유사도 MS_b를 측정하는데 있어 형태정보 i를 M_Q에 포함될 때 Q와 D사이의 MS_b의 값은 다음식과 같이 결정된다.

$$MS_b(Q,D)=\sum_{\forall i\in M_Q}(M_{Qi}, M_{Di}) \quad (5)$$

$MS_b(Q, D)$값이 작을 수록 형태는 더 유사하고, 만약 $MS_b(Q, D)$이 크면 두 영역간에는 형태 유사도가 없다는 것을 뜻하며, M_{Qi}, M_{Di}는 질의 및 참조영상의 5개의 형태정보를 나타낸다. 모든 형태정보 쌍들에

대해 계산되어진 유사값들은 similarity matrix(SM)이라 불리는 M×N 행렬에 저장된다.

3.3.4 가중치 유사도 계산 함수

칼라 공간 및 형태검색 방법에 있어서 본 논문에서는 칼라, 공간, 형태의 각 특징에 가중치(weight)를 지정할 수 있게 하였다. 응용에 따라서 객체의 형태에 더 높은 비중을 줄 수도 있으며, 객체의 칼라에 더 높은 비중을 줄 수도 있다. 이는 유사한 칼라를 갖는 물체가 가운데 위치해 있거나, 비슷한 위치에 유사한 객체를 가지는 영상을 검색할 때에 유용하다. 가중치를 부여한 유사도 SIM_W 측정은 다음식과 같다. 다음 그림 7은 가중치에 의한 영상 유사도 검색 방법을 설명하고 있다.
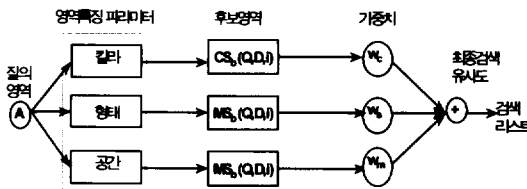
그림 7. 가중치에 의한 유사도 검색

$$SIM_W(Q,\ D) = \sum_{\forall i \in C_Q} CS_b(Q, D, i) * W_C + \sum_{\forall i \in S_Q} SS_b(Q, D, i) * W_S + \sum_{\forall i \in M_Q} MS_b(Q, D, i) * W_M \quad (6)$$

여기서, C_Q와 S_Q및 M_Q는 Q영상의 칼라, 공간, 형태를 나타내는 집합이고 W_C, W_S, $W_M(< 1)$은 영상의 칼라, 공간 및 형태특징에 주어지는 가중치이다.

4. 시스템 구현 및 실험

4.1 시스템 구현

본 논문에서 구현한 내용기반 영상검색 시스템은 Win 95/Win NT 운영체제로하는 PC Pentium상에서 구현하였으며 데이터베이스는 Apollo pro 4.5를 사용하고 개발S/W 언어로서 검색엔진을 Visual C++ 6.0으로 동적 라이브러리(DLL)화하여 델파이 4.0으로 구현된 텍스트 검색 엔진에 탑재한 형태로 구현하였다. 전체적인 시스템은 다섯가지 주요 모듈 즉, 사용자 인터페이스, 텍스트 검색, 영상 검색, 프리젠테이션, 데이터베이스 색인 모듈로 이루어져 있다.

영상 데이터베이스는 테스트를 위해 사용되는 3,000개의 영상 집합으로 이루어져 있으며, 여러개의 카탈로그로 나누어지는데 주요 카탈로그는 동물, 미술, 컴퓨터, 음식, 자연, 과학, 여행 등을 포함하고 있다. 본 논문에서는 검색결과 프리젠테이션에 사용되는 thumbnail 영상들을 120×N 픽셀 혹은 N×115 픽셀의 크기로 축소하였다.
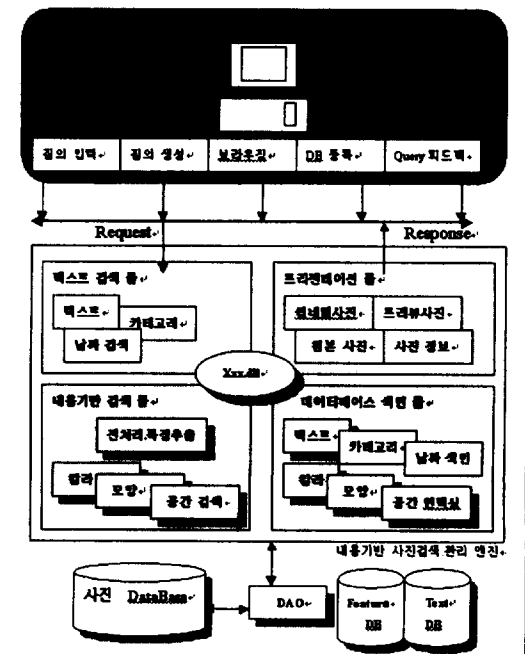
그림 8. 영상 검색을 위한 시스템 구성도

4.2 검색 실험 및 평가

수만개 이상의 영상이 모인 데이타베이스에서 질의와 관계된 모든 영상들을 검색하는 것은 속도면에서 바람직하지 않다. 본 실험에서는 유사도 측정에 의해 랭크된 상위 60개의 영상들만 평가한다. 따라서 모든 관련된 영상들이 검색되지 않기 때문에 일반적으로 사용되는 Recall-Precision 평가방법은 적당하지 않다. 대신 각각의 질의에 대해 정규화된 정확도(P_n)와 정규화된 재현율(R_n)를 계산한다. P_n과 R_n의 식은 다음과 같으며 여기서 N은 총 영상들의 수이고 R은 총영상들 중 관련된 영상들의 수이다. 그리고 rank i는 시스템에 의해 검색되어지는 관련된 영상 i의 랭크된 순서를 나타낸다.

$$P_n = 1 - \frac{\sum_{i=1}^{R}(\log\ rank_i - \log\ i)}{\log \frac{N!}{(N-R)!R!}} \qquad (7)$$

$$R_n = 1 - \frac{\sum_{i=1}^{R}(rank_i - i)}{(N-R)R} \qquad (8)$$

계산에서 N은 3,000으로 정하고 상위 60개에 배열되지 않은 관련된 영상들은 그들의 배열을 60과 3,000 사이의 중간에서 가까운 값으로 정하였다.

본 논문에서 영상 특징 데이타베이스 구축에 사용된 참조 영상에 대한 범주는 다음 표 1과 같이 3,000여 개이다. 본 논문에서는 다양한 종류의 영상에 대하여 특정 형태의 영상 검색 시스템 효율과 성능을 알아보기 위해서는 영상들을 두 개의 카테고리로 분류하였다. 즉, 객체가 구분되지 않은 영상들과 구별되는 객체를 갖는 영상들이다. 예를 들면 첫 번째 카테고리는 해안가나 설원 풍경과 같은 영상들이고 두 번째 카테고리는 붉은 꽃, 공모양등 구별되는 객체를 갖는 영상들이다. 질의영상은 첫번째 카테고리로부터 Q_1에서 Q_5까지 5개, 두번째 카테고리로부터 Q_6에서 Q_8까지 3개등 총 8개의 영상들을 무작위로 선택하였다.

다음 그림 9와 그림 10은 첫 번째 카테고리에 속한 영상들에 대하여 질의한 검색결과들이다. 그림9는 선택한 샘플영상에 질의에 대한 칼라 검색결과이고 그림10은 선택한 샘플영상에 질의에 대한 공간 검색 결과이다. 다음 그림 10과 그림 11은 두 번째 카테고리에 속한 영상들에 대하여 질의한 검색결과들이다. 그림 11은 첫번째 선택한 샘플영상에 질의에 대한 형태 검색결과를 보이고 있다. 그림 12는 첫번째 샘플영상에 질의에 대한 칼라-공간 및 형태정보의 가중치 검색결과를 보이고 있다.

영상검색 실험 결과를 표 2와 표 3에 요약하였다.

표 1. 참조영상의 범주

범주	대분류	개수	범주	대분류	개수
(1)	과 학	300	(5)	식 물	600
(2)	동 물	400	(6)	자 연	900
(3)	산 업	400	(7)	문화예술	100
(4)	스포츠	300	합 계		3,000
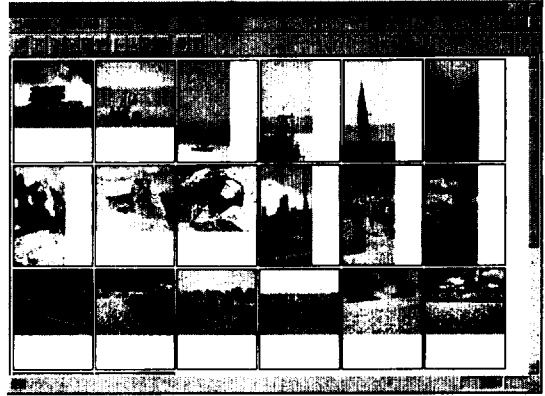
그림 9. 샘플영상 질의에 대한 칼라검색 결과
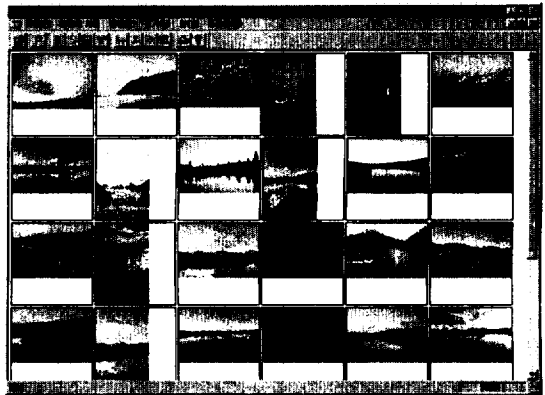
그림 10. 샘플영상 질의에 대한 공간 검색 결과
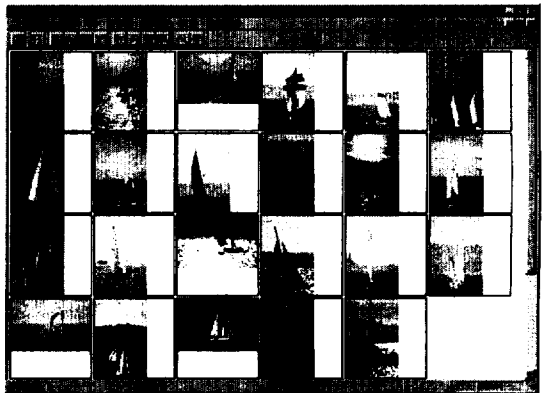
그림 11. 샘플영상 질의에 대한 칼라검색 결과

표 2에서 대표칼라의 수를 8로 고정시키고 임계치값을 0.25~0.55까지 변화시켜 가면서 실험한 결과, 임계치값이 0.4일 때 정규화된 Precision 및 정규화된
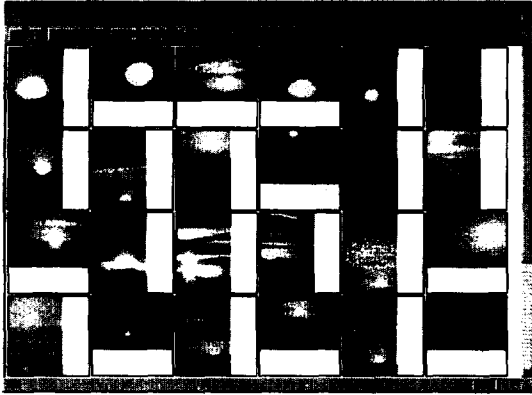
그림 12. 샘플영상 질의에 대한 칼라,공간, 형태 검색 결과

Recall이 각각 평균 0.60, 0.73으로 가장 적합하였다. Q1~Q5는 배경이 복잡하여 객체가 구별되지 않는 질의 영상들을 사용한 경우이고, Q6~Q8은 배경과 객체가 구별되는 질의영상을 사용한 경우이다. 따라서, Q6의 경우 정규화된 Precision 및 정규화된 Recall이 각각 0.74, 0.85로 가장 우수한 검색효율을 보였다.

표 3에서는 시스템의 효과적인 칼라검색에 있어 효율을 분석하였다. 표 2를 토대로 가장 효율이 좋은 대표칼라 설정 임계치는 0.4로 고정하였고, 대표 칼라수를 3, 4, 5개까지 변화시켜가면서 실험하였다. 표 3에서 보는바와 같이 대표칼라를 4개 선택하여 실험한 결과 정규화된 Precision 및 정규화된 Recall이 각각 평균 0.60, 0.73으로 가장 좋은 효율을 보였고, 질의 영상중 Q6의 경우가 정규화된 Precision 및 정규화된 Recall이 각각 평균 0.74, 0.85로 여러개의 질의 영상중 가장 좋은 검색 결과로 나타났다.

본 논문에서 제시한 유사도 측정은 객체의 칼라,

표 2. 임계치 변화에 따른 정규화된 정확도와 재현율

	대표칼라 설정을 위한 임계치											
	정규화된 Precision						정규화된 Recall					
	0.30	0.35	0.40	0.45	0.50	0.56	0.30	0.35	0.40	0.45	0.50	0.55
Q1	0.57	0.60	0.63	0.59	0.56	0.51	0.74	0.71	0.75	0.73	0.70	0.70
Q2	0.32	0.40	0.43	0.40	0.35	0.33	0.57	0.66	0.64	0.64	0.63	0.62
Q3	0.34	0.42	0.46	0.43	0.42	0.40	0.58	0.57	0.58	0.56	0.54	0.53
Q4	0.52	0.57	0.60	0.60	0.58	0.54	0.68	0.69	0.71	0.71	0.68	0.67
Q5	0.51	0.58	0.62	0.61	0.58	0.50	0.78	0.79	0.70	0.79	0.78	0.74
Q6	0.70	0.72	0.74	0.73	0.71	0.68	0.84	0.83	0.85	0.83	0.81	0.80
Q7	0.62	0.65	0.63	0.60	0.56	0.52	0.82	0.81	0.79	0.67	0.64	0.66
Q8	0.56	0.58	0.65	0.64	0.60	0.56	0.77	0.76	0.80	0.83	0.8	0.73
Avg	0.52	0.57	0.60	0.58	0.55	0.51	0.72	0.73	0.73	0.72	0.70	0.68

표 3. 칼라수 변화에 따른 정규화된 정확도와 재현율

	칼라수 변화 C					
	정규화된 Precision			정규화된 Recall		
	3	4	5	3	4	5
Q1	0.60	0.63	0.62	0.72	0.75	0.73
Q2	0.41	0.43	0.44	0.62	0.64	0.64
Q3	0.44	0.46	0.46	0.55	0.58	0.57
Q4	0.60	0.60	0.58	0.71	0.71	0.70
Q5	0.60	0.62	0.59	0.67	0.70	0.69
Q6	0.71	0.74	0.73	0.84	0.85	0.84
Q7	0.60	0.63	0.61	0.76	0.79	0.78
Q8	0.62	0.65	0.63	0.78	0.80	0.80
Avg	0.57	0.60	0.58	0.71	0.73	0.72

공간 및 형태 특징에 대한 가중치를 부여할 수 있다. 구별되는 객체를 가지지 않는 영상들은 이러한 유사도 측정을 만족할 수 없기 때문에 가중치 검색실험은 Q_6~Q_8의 질의영상들로 제한하였다. 각각의 질의는 정규화된 정확도 및 재현율에 의한 실험결과는 표 4에 나타냈다. 표 4는 객체가 명확히 구별되는 질의영상에 대한 가중치 검색결과를 분석한 표이다. 칼라-공간, 칼라-형태, 공간-형태, 칼라-공간-형태와 같은 방법으로 가중치 검색 실험을 하였고, 그중 칼라-공간-형태를 조합한 검색효율이 정규화된 Precision 및 정규화된 Recall이 각각 평균 0.71, 0.86으로 가장 우수하였다.

기존의 칼라 히스토그램 방법은 검색소요되는 시간이 평균적으로 30초 이상되기 때문에 성능이 매우 떨어진다. 본 논문에서 각 질의에 대해 검색에 소요되는 시간이 약 6~7초 정도되는 소요되었다. 이것은

표 4. 가중치 검색 비교

	정규화된 Precision			
	칼라-공간	칼라-형태	공간-형태	칼라-공간-형태
Q6	0.74	0.76	0.72	0.77
Q7	0.63	0.65	0.62	0.67
Q8	0.65	0.67	0.63	0.69
Avg	0.67	0.69	0.66	0.71

	정규화된 Recall			
	칼라-공간	칼라-형태	공간-형태	칼라-공간-형태
Q6	0.85	0.87	0.83	0.89
Q7	0.79	0.81	0.77	0.83
Q8	0.80	0.83	0.78	0.85
Avg	0.81	0.84	0.79	0.86

참고문헌[6]에서 권장하는 10초 한계를 만족하는 시간이다. 그러나, 이와 반대로 가중치를 사용한 검색 소요 시간은 약 40초 정도이다. 이는 영상의 칼라, 공간, 형태정보를 쌍방향으로 교차하여 비교하기 때문이다.

4.3 본 시스템의 고찰

표 5는 이런 3가지 영상들의 검색실험 결과를 보이고 있는데, 정규화된 정확도와 재현율이 기존 히스토그램 검색 방법보다 낮다. 실험조건은 표2와 동일하게 대표 칼라수를 3, 4, 5개까지 변화시켜가면서 실험하였으며, 대표칼라를 5개 선택하여 실험한 결과 정규화된 Precision 및 정규화된 Recall이 각각 평균 0.56, 0.66으로 점점 좋은 검색 효율을 보였는데, 이는 선택된 몇 개의 대표 칼라가 전체 영상을 대표하지 못하므로 기존의 히스토그램 방법처럼 칼라수를 많이 선택하면 할수록 검색 효율이 높아짐을 알 수 있다.

다음 표 6는 본 논문의 효율성을 입증 하기위해 타 연구와의 실험결과를 비교한 것이다. 일본 도쿄대학의 GOLS[21]는 170 여개의 산,물,하늘 잔디, 일몰과 같은 화상에 대하여 장면묘사(Scene description keyword)[22]를 사용하여 화상검색실험을 하여 precision과 recall 이 각각 0.645, 0.854의 검색을 얻었고 일본 NTT 연구소에서는 160개의 꽃과 100여개의 애니메이션영화의 클립화상에 대해 FCI(Focused Color Intersection)방법을 사용하여 화상검색을 하여 precision과 recall을 0.69와 0.85을 얻었다. 본 연구와 기존연구와의 연구는 검색에 적용하는 화상의 종류나 검색성분 및 실험환경이 다르기 때문에 절대비교는 불가능하지만 본 연구의 실험결과가 비교적 우수하였다.

표 5. 칼라가 산재한 영상에 대한 검색 비교

	칼라수 변화 C					
	정규화된 Precision			정규화된 Recall		
	3	4	5	3	4	5
Q9	0.50	0.53	0.55	0.65	0.69	0.72
Q10	0.37	0.40	0.43	0.60	0.62	0.63
Q11	0.40	0.43	0.45	0.51	0.55	0.57
Q12	0.51	0.55	0.56	0.66	0.68	0.69
Q13	0.56	0.60	0.63	0.63	0.67	0.68
Q14	0.52	0.57	0.58	0.60	0.63	0.65
Q15	0.54	0.61	0.64	0.52	0.59	0.61
Q16	0.55	0.62	0.66	0.63	0.70	0.74
Avg	0.49	0.54	0.56	0.60	0.64	0.66

표 6. 타 연구와 실험결과 비교

특징성분 파라메타	GOLS	NTT Lab	본 연구
총화상수	170	260	3,000
검색성분	장면묘사단어	FCI	칼라-공간-형태
Precision	0.64	0.69	0.71
Recall	0.85	0.85	0.86

5. 결 론

본 논문에서는 영상의 칼라-공간 및 형태 정보를 추출하고 검색하는 새로운 방법을 제시하였다. 제시한 기술을 지원하는 본 영상검색 프로토타입 시스템으로 3,000개의 영상들을 검색 테스트하고 실험하였다. 본 논문에서는 알고리즘을 히스토그램를 기반으로 한 기술과 집약된 칼라-공간 기술과 비교하였으며 그 결과는 효율적인 면에서 우리의 방법이 집약된 칼라-공간 기술보다 성능이 우수하고 히스토그램를 기반으로 한 기술 보다 평균적으로 정규화된 정확도와 정규화된 재현율에서 월등히 향상되었음을 입증하였고, 또한 두 기술보다 더 빠른 검색이 가능함을 보였다. 그러나 제시한 기술로는 주된 칼라가 없는 영상들과 주된 칼라들이 같이 모여 있지 않는 영상들을 효과적으로 다룰 수 없는 문제가 있으므로 이를 보완하는 연구와, 칼라 영역, 공간, 형태 정보에 대한 색인 시간을 감소시키는 문제에 대해 연구해야 한다.

참 고 문 헌

[1] A.One and T.Satou. Image database with fully automated keyword extraction function by integration state transition model and vsdl. Proc. of 6th Symposium of Runctional Graphics Information System, pp. 7-10, 1995.

[2] H. Kasahara and T.Kishimoto. Pictorial database navigation. IEICEJ Technical Report IE88-46, pp. 71-78, 1988.

[3] H.Nagata and K.Aiba. Image retrieval method based on memory and contents of an image. IPSJ National Conf. 1L-6, pp. 101-102, 1991.

[4] J. Yamance and M.Sakauchi. A construction of a new image database system which realize fully automated image keyword extraction . IEICEJ Transition on Information and Systems, 10(E76-D), pp. 1211-1233, 1993.
[5] M.Flickner. Query by image and video contemt : the qbic system. IEEE computer, pp. 23-22, 1995.
[6] M.Minoh and H..Okazaki. A retrieval method of image database using attributive features of objects-in case of a mountain in the scene image as an example-. IPSJ Journal, 32(4), pp. 513-522, 1991
[7] M.Mukumot and M.Minoh. Scenary image retrieval using index given by image recognition technique. Proc. of 141st IEEE workshop 94-02-01, pp. 1-4, 1994.
[8] M.Sakauchi Imiage retrieval technoloy. IEEE Journal, 71(9), pp. 911-914, 1988.
[9] M.Sakauchi Information access for multimedia. Report of the IIS the Univ Tokyo , 44(11), pp. 2-6, 1992
[10] M.Sakauchi. Database vision and image retrieval. IEEE Multimedia, 1(1), pp. 79-81, 1995.
[11] M.Shicata. An associative retrieval method for an image database. IEICEJ Technical Report PRU88-138, pp. 3-30, 1988.
[12] S.Yamano and Y.Yaginuma. Experiment and evaluation on image database construction using vector expression of hue and shape. IPSJ National Conf. No. 4G-7, pp. 129-130, 1995
[13] T,Kato and H. Simogaki. Trademark: multmedia image database system with intelligent human interface IEICEJ Trans.(D-II), J72-D-II(4), pp. 535-544, 1989.
[14] T.Kato and T.Kurita. Electronic art museum: Full color image database with visual interaction on color and sketch. IEICEJ Technical Report IE88-118, pp. 31-38, 1988.
[15] T.Satoh and M.Sakauchi. Video scene acquisition on live hypermedia. Proc. 2nd IEEE International Conference on Multimedia Computing and Systems, pp. 175-181, 1995.
[16] T.Satou and M.Sakauchi. A new type hypermedia platform for industrial applications. Proc. of IECON'93, 1993.
[17] T.Takahasi and N.Shima. An image database retrieval system using spatial relationships. IEICEJ Technical Report PRU89-80, pp. 23-28, 1989.
[18] V.E.Ogle and M.Stonbraker. Chabot:retrieval from a relational database of images. IEEE compuer, pp. 40-48, 1995.
[19] Y.Gong and H.Zhang. An imag database system with-content captureing and fast image indexing abilities. Proc. 1st IEEE Inter. Conf. on Multimedia Computing and Systems, pp. 121-130, 1994.
[20] Y.Yaginuma and M.Sakauchi. Multi-purpose interface for still/moving image retrieval. Proc. of the SPIE International Conference of Image Processing and Its Applications, pp. 260-267, 1994.
[21] Atsushio Ono et al., A flexible contented-based image retrieval system with combined scene description keyword," in Int. conf. on Multimedia computing and system, Hiroshima Japan pp. 201-208 June 1996.
[22] V Vinoid et al., Focused color intersection with efficient searching for object detection and image retrieval," Int. conf on Hiroshima Japan pp. 229-233 June 1996.

김 성 동

1981년 광운대학교 응용전자공학과 졸업
1984년 아주대학교 대학원 전자공학 석사
1991년 뉴욕시립대학교 대학원 전산과 석사
1994년 미국 스티븐슨 공대 박사수료
1998년 광운대학교 컴퓨터공학과 박사수료
1995년~현재 계원조형예술대학 정보통신과 교수
관심분야 : Object recognition and tracking, digital image processing Structural and Syntactic pattern recognition

최 기 호

1973년 한양대학교 전자공학과 졸업
1977년 한양대학교 대학원 전자공학과 석사
1987년 한양대학교 대학원 전자공학과 박사
1977~1979년 한국과학기술원 연구소 연구원
1979 ~ 현재 광운대학교 컴퓨터공학과 교수
1989~1990년 University of Michigan Visiting Scholar
관심분야 : 멀티미디어 정보검색, Object recognition and tracking